数理解析研究所講究録
265巻 1976年 93-104

L^1空間における最良近似と統計への応用

大阪市大　理学部　工藤弘吉

1. 序　確率変数Xのメジアン$\mu(X)$はXの平均偏差（または平均絶対誤差）$E[|X-a|]$を最小にするようなaの値であることは周知の事実である.（例えばCramer [3, p.178]）.一方条件付メジアンの定義も例えばLoeve [11, p.385] などにも厳密に与えられているのであるから, それが条件付平均偏差を最小にするものであることは当り前のことと了解されている. ところが函数解析の立場からは, L^1空間での部分空間における最良近似が必らずしも存在しないという事柄からいって, 上のことは自明だとは云えないということである. そこでこの事に関して新谷・安藤論文[16]の結果を中心にして第2節で述べてみたい.

筆者がこのような問題に関係するようになったのは, 部分σ代数の収束の問題, 特に下の[C-1]と[B-2]との同等性の証明に関してである. 部分σ代数の収束についてはかなり多様な

定義が可能であって，それらの定義の関係は複雑であり，この収束については第3節で述べる．

最後の節では部分σ代数の収束の問題の発端となった漸近十分統計量の問題について報告する．最近の数理統計学では統計量と部分σ代数とはほとんど同義語として取扱っても差支ない．特に統計量の十分性の問題は統計量の関数としての性格よりは部分σ代数としての性格に関するものであって，その収束を論ずる場合はどうしても部分σ代数の収束の意味を考えざるを得ないことになる．漸近十分統計量の概念は筆者が導入したものであるが，それが部分σ代数の収束の概念を通じて在来の十分性と密接な関係にあることがわかってきた．

2. 最良近似　確率空間$(\Omega, \mathcal{A}, P)$は，集合Ω，その部分集合のσ代数$\mathcal{A}$および$\mathcal{A}$上で定義された確率測度Pからなりたつ．以下$\mathcal{A}$の部分σ代数$\mathcal{B}$はすべて$(\mathcal{A}, P)$完備（$A \in \mathcal{A}$, $B \in \mathcal{B}$ かつ $P(A \triangle B)=0$ なら$A \in \mathcal{B}$となること）であるとする．Pの部分σ代数$\mathcal{B}$への制限もまた同じ記号Pであらわすことにする．そのとき$(\Omega, \mathcal{B}, P)$もまた一つの確率空間であり，そこでp乗可積分な実数値函数の全体を$L^p(\mathcal{B})$であらわそう．そうすると，$L^1(\mathcal{B})$は$L^1(\mathcal{A})$の部分空間である．

$f\ (\in L^1(\mathcal{A}))$の平均を$E[f]=\int f\,dP$（$P$を明示する必要があるときは$E_P[f]$）であらわす．$f\ (\in L^1(\mathcal{A}))$の

$\mathcal{B}$ ($\subset \mathcal{A}$) が与えられた條件付平均 $E^{\mathcal{B}}[f]$ (または $E_P^{\mathcal{B}}[f]$) は，$\forall B(\in \mathcal{B})$: $E[f\cdot 1_B] = E[1_B \cdot g]$ をみたす $g(\in L^1(\mathcal{B}))$ をあらわす．A ($\in \mathcal{A}$) の $\mathcal{B}$ が与えられた条件付確率 $P^{\mathcal{B}}(A)$ は $E^{\mathcal{B}}[1_A]$ のことである．$f(\in L^1(\mathcal{A}))$ の$\mathcal{B}$が与えられた条件付メジアンは $P^{\mathcal{B}}\{f<g\} \leqq \frac{1}{2} \leqq P^{\mathcal{B}}\{f \leqq g\}$ なる $g(\in L^1(\mathcal{B}))$ である．どんな f ($\in L^1(\mathcal{A})$) に対しても，その $\mathcal{B}$ が与えられた条件付メジアンは存在するが一意的とは限らない．ただし条件付メジアンには最大のもの $M^{\mathcal{B}}[f]$，最小のもの $m^{\mathcal{B}}[f]$ が存在して，$m^{\mathcal{B}}[f] \leqq g \leqq M^{\mathcal{B}}[f]$ なる $g \in L^1(\mathcal{B})$ はすべて $\mathcal{B}$ の与えられた条件付メジアンである．$f(\in L^1(\mathcal{A}))$，$\mathcal{B}$ ($\subset \mathcal{A}$) が与えられたとき，$L^1(\mathcal{A})$ のノルムの意味で f に最も近い g ($\in L^1(\mathcal{B})$) を f の$\mathcal{B}$可測最良近似という．どんな f ($\in L^1(\mathcal{A})$) に対しても $\mathcal{B}$ 可測最良近似と $\mathcal{B}$ が与えられた条件付メジアンとは完全に一致する．([16, 定理2], [9, Appendix 定理A.1])．

f ($\in L^1(\mathcal{A})$) からその $\mathcal{B}$ 可測最良近似までの距離は f から $L^1(\mathcal{B})$ までの最短距離 $e_P(f, \mathcal{B})$ (P が前後の脈絡から明らかなときは P を略して $e(f, \mathcal{B})$ とかく．第3節までは P を略す) であるが，それは

(2.1) $$e(f, \mathcal{B}) = \int_{-\infty}^{\infty} \left\{ \frac{1}{2} - \int \left| \frac{1}{2} - P^{\mathcal{B}}\{f \leqq t\} \right| dP \right\} dt$$

で表わされる．特に $f \equiv 1_A$ ($A \in \mathcal{A}$) のときは，それは

$$(2.2)\qquad d(A,\mathcal{B})=\min P(A\triangle B)=\frac{1}{2}-\int\left|\frac{1}{2}-P^{\mathcal{B}}(A)\right|dP$$

と一致する.（［16，系4］，［8；Remark 2.1］）. この量は距離$P(A\triangle B)$によって$\mathcal{A}$を距離空間とみなしたとき，Aから閉集合$\mathcal{B}$までの最短距離である. ここで$\triangle$は集合の対称差を表わす.

$\{g_n\}$（$g_n\in L^1(\mathcal{B})$）が $\lim_{n\to\infty}\|g_n-f\|_1=e(f,\mathcal{B})$ をみたすとき，$\{g_n\}$は相対弱コンパクトであって，その弱極限はfの$\mathcal{B}$可測最良近似である（［16，定理5］）

この定理から直ちに，上記の$\{g_n\}$は$\|g_n\vee M^{\mathcal{B}}[f]-M^{\mathcal{B}}[f]\|_1\to 0$，$\|g_n\wedge m^{\mathcal{B}}[f]-m^{\mathcal{B}}[f]\|_1\to 0$ $(n\to\infty)$ をみたす（［16，系6］）.

なお条件付メジアンの性質についてはTomkins［17］をみよ.

3. 〜代数の収束　$\{\mathcal{B}_n\}$を$\mathcal{A}$の部分σ代数$\mathcal{B}_n$の列とする. まずはじめに$\{\mathcal{B}_n\}$の収束の定義を筆者の知る限り列挙すれば次の通りである.

［A］ 集合論的定義：$\cap$は集合論的な共通部分，$\vee$は"…を含む最小のσ代数"をあらわす記号とする.

(1) $(s)\text{-}\overline{\lim}\,\mathcal{B}_n=\bigcap_{m=1}^{\infty}\bigvee_{n=m}^{\infty}\mathcal{B}_n$，$(s)\text{-}\underline{\lim}\,\mathcal{B}_n=\bigvee_{m=1}^{\infty}\bigcap_{n=m}^{\infty}\mathcal{B}_n$

として，$(s)\text{-}\overline{\lim}\,\mathcal{B}_n=(s)\text{-}\underline{\lim}\,\mathcal{B}_n$（$=\mathcal{B}_0$とおく）のとき

$\mathcal{B}_0=(s)\text{-}\lim\mathcal{B}_n$ とする.

もちろん$\{\mathcal{B}_n\}$が単調なら（$\mathcal{B}_n\subset\mathcal{B}_{n+1}$または$\mathcal{B}_{n+1}\subset\mathcal{B}_n$なら）$(s)\text{-}\lim_{n\to\infty}\mathcal{B}_n$は存在する. この収束は確率測度に関係ない. なお

Nghien [14]は上の収束より弱いものを定義しているが，これも確率測度Pに関係ないので，われ〳〵の目的にあまり影響がないから省略する．

［B］<u>幾何学的定義</u>　距離$P(A \triangle B)$によってσ代数$\mathcal{A}$を距離空間とみなすとき，$\mathcal{A}$の部分σ代数は$\mathcal{A}$の閉部分集合であり，$d(A, \mathcal{B})$はAから$\mathcal{B}$までの距離である．

（1）Boylan［2］の同程度収束．2つの部分σ代数$\mathcal{B}$，$\mathcal{B}'$に対して　$D(\mathcal{B}, \mathcal{B}') = \sup_{B' \in \mathcal{B}'} d(B', \mathcal{B}) + \sup_{B \in \mathcal{B}} d(B, \mathcal{B}')$　とおくと，部分σ代数全体$\mathbf{B}$はDを距離とする距離空間となる．この距離空間は完備である．距離Dによる収束をB収束ということにする．

Rogge［15］は上のDを少し修正したが，収束に関する限りでは特記する必要はない．

（2）$\forall A (\in \mathcal{A}) : d(A, \mathcal{B}_n) \to d(A, \mathcal{B}_0)$　のとき，$\{\mathcal{B}_n\}$が$\mathcal{B}_0$にd弱収束するという．

（3）岩田弱収束：$(i)\text{-}\underline{\lim}\, \mathcal{B}_n = \{ A (\in \mathcal{A}) : \lim_{n \to \infty} d(A, \mathcal{B}_n) = 0 \}$，$(i)\text{-}\overline{\lim}\, \mathcal{B}_n = \bigvee \{ A (\in \mathcal{A}) : \underline{\lim}_{n \to \infty} d(A, \mathcal{B}_n) = 0 \}$とおき，$(i)\text{-}\underline{\lim}\, \mathcal{B}_n = (i)\text{-}\overline{\lim}\, \mathcal{B}_n \ (= \mathcal{B}_0)$のとき，$\mathcal{B}_n$は$\mathcal{B}_0$に岩田弱収束という（岩田［5］）．

ここで$(i)\text{-}\underline{\lim}\, \mathcal{B}_n$はσ代数であることは容易に示される．

［C］<u>函数解析的定義</u>　$L^p(\mathcal{A})$の閉部分空間の列$\{L^p(\mathcal{B}_n)\}$

の収束によって$\{\mathcal{B}_n\}$の収束を定義する.

（1）　$\forall f(\in L^p(\mathcal{A}))$: $e(f,\mathcal{B}_n)\to e(f,\mathcal{B}_0)$であるとき, $\mathcal{B}_n$は$\mathcal{B}_0$にe弱収束するという

（2）　Becker弱収束. σ代数$\mathcal{B}$, $\mathcal{B}'$, $g(\in L^\infty(\mathcal{A}))$に対して, $D(\mathcal{B},\mathcal{B}'|g)=\sup_{f\in L^1(\mathcal{B})}\inf_{f'\in L^1(\mathcal{B}')}|\int(f-f')g\,dP|+\sup_{f'\in L^1(\mathcal{B}')}\inf_{f\in L^1(\mathcal{B})}|\int(f-f')g\,dP|$とおき, $\forall g(\in L^\infty(\mathcal{A}))$: $D(\mathcal{B}_n,\mathcal{B}_0|g)\to 0$ となるとき$\mathcal{B}_n$は$\mathcal{B}_0$にBecker弱収束という（Becker [1]）. このDは一様構造を与えるがそれは完備である. なお上記の収束はBeckerの与えた位相をやゝ変形して収束の形で述べたものである.

［D］ <u>条件付平均の収束による定義</u>　確率変数の収束の種類に伴って多くの種類の収束が定義できる.

（1）　$\mathcal{B}_n$が$\mathcal{B}_0$にa.e.収束するとは, $\forall f(\in L^p(\mathcal{A}))$: $E^{\mathcal{B}_n}[f]\to E^{\mathcal{B}_0}[f]$, a.e.　マルチンゲールの理論によると, $\{\mathcal{B}_n\}$が単調列であれば, この収束は集合論的収束［A-1］に一致する.

（2）　$\mathcal{B}_n$が$\mathcal{B}_0$に確率収束するとは, $\forall f(\in L^p(\mathcal{A}))$: $E^{\mathcal{B}_n}[f]\to E^{\mathcal{B}_0}[f]$ in prob.

（3）　$\mathcal{B}_n$が$\mathcal{B}_0$に強収束（またはL^p収束）するとは, $\forall f(\in L^p(\mathcal{A}))$: $\|E^{\mathcal{B}_n}[f]-E^{\mathcal{B}_0}[f]\|_p\to 0$, ただし$1\leqq p\leqq\infty$とする.

（4）　$\mathcal{B}_n$が$\mathcal{B}_0$に弱収束するとは, $\forall f(\in L^p(\mathcal{A}))$, $\forall g(\in L^q(\mathcal{A}))$: $\int E^{\mathcal{B}_n}[f]g\,dP\to\int E^{\mathcal{B}}[f]g\,dP$, ただし$p\in[1,\infty)$, $\frac{1}{p}+\frac{1}{q}=1$.

(2), (3)はNeveu [12; p.124] に載っているものであり，この意味の収束では $\mathbf{B}$ は完備空間である．(4)はBecker [1] で僅かにふれられているにすぎない収束概念であるが，もちろんBecker弱収束とは異なる概念であり，ここでの収束の意味では $\mathbf{B}$ は完備ではないことに注意されたい．

条件付平均を用いることによってさらにもう1つの収束を定義できる．一般的には確率変数列のノルムが収束するからといって，その極限が定まるということはないのであるが，条件付平均の列の場合には，そのノルムが収束するとき，部分σ代数列の極限が定まるのである．それは条件付平均のノルムがσ代数の大きさに依存するという事実に基くのである．

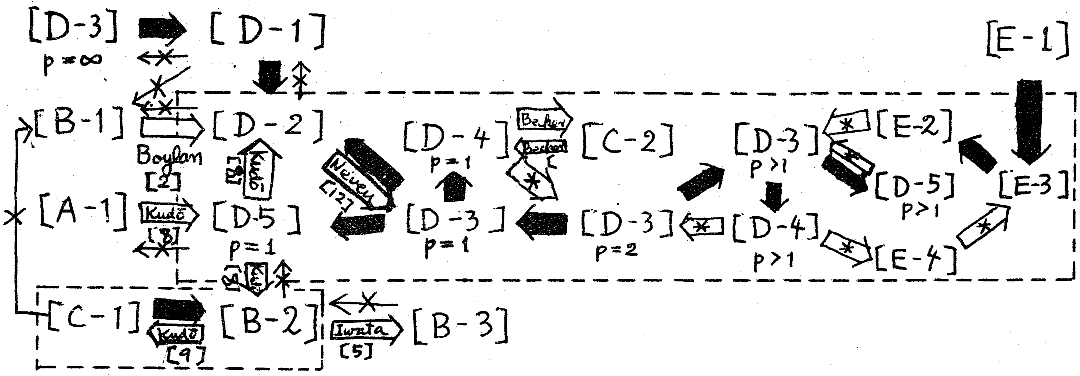
(5) "$\forall f\,(\in L^p(\mathcal{A}))$: $\overline{\lim_{n\to\infty}}\,\|E^{\mathcal{B}_n}[f]\|_p \leqq \|E^{\mathcal{B}}[f]\|_p$ $\left(\underline{\lim_{n\to\infty}}\,\|E^{\mathcal{B}_n}[f]\|_p \geqq \|E^{\mathcal{B}}[f]\|_p\right)$" をみたすσ代数 $\mathcal{B}$ の中で最小（最大）のものが存在する．この最小（最大）のものを $(p)\text{-}\overline{\lim}\,\mathcal{B}_n$ $\left((p)\text{-}\underline{\lim}\,\mathcal{B}_n\right)$ とおく．$(p)\text{-}\underline{\lim}\,\mathcal{B}_n = (p)\text{-}\overline{\lim}\,\mathcal{B}_n$ がなりたつとき，$\mathcal{B}_0 = (p)\text{-}\lim\,\mathcal{B}_n$ と定義する．（$p=1$ についてはKudō [8]，$p>1$ については安藤氏よりの私信）．なお岩田[5]によれば，$(s)\text{-}\underline{\lim}\,\mathcal{B}_n \leqq (i)\text{-}\underline{\lim}\,\mathcal{B}_n = (p)\text{-}\underline{\lim}\,\mathcal{B}_n \leqq (p)\text{-}\overline{\lim}\,\mathcal{B}_n \leqq (i)\text{-}\overline{\lim}\,\mathcal{B}_n \leqq (s)\text{-}\overline{\lim}\,\mathcal{B}_n$ ということである

さらに，$p>1$ のときは，条件付平均 $E^{\mathcal{B}}$ の代りに最良近似（＝条件付メジアン，予測）$\mu^{\mathcal{B}}$ を用いることができる．

[E] 最良近似の収束による定義　$p=1$ のときは $\mu^{\mathcal{B}}f$ が一意的に定まらない（2節参照）から，$p>1$ とする．そのとき，[D] の $E^{\mathcal{B}}$ の代りに $\mu^{\mathcal{B}}$ を用いることによって全く同様に $\mathcal{B}_n$ の収束が定まるから，それらを [D] における番号に対応して [E-1] —[E-4] と名づける．

以上述べた5種類の定義の間には同等なものがかなりあって単純にはゆかない．下図はそれらの関係を示す．

（σ代数の収束概念関連図）



この図で ➡ は特別な証明を必要としないものであり，著者名[2] は著者名と論文を示す．[※]⇨ は安藤氏より筆者への私信による．↛ は反例が示めされているもの，┆┄┆は同等なもの．

4. 統計への応用　$\mathcal{P}$ を $(\Omega, \mathcal{A})$ 上の確率測度の集合とする．組 $(\Omega, \mathcal{A}, \mathcal{P})$ を統計的実験または統計空間 —— これは筆者の命名によるもの —— という．$\mathcal{A}$ の部分σ代数 $\mathcal{B}$ に対する $f\ (\in L^1(\mathcal{A}))$ の条件付平均 $E_P^{\mathcal{B}}[f]$ が，どんな f に対しても，$P\ (\in \mathcal{P})$ に無関係に定まるとき —— いいかえれば，

$\forall f(\in L^1(\mathcal{A}));\exists g(\in L^1(\mathcal{B})):\forall P(\in\mathcal{P}),\forall B(\in\mathcal{B}):\int_B f\,dP = \int_B g\,dP$ であるとき——$\mathcal{B}$は$\mathcal{P}$に対して十分であるという. ある確率測度λに関して$\mathcal{P}$のどの要素も絶対連続であるとき, λを$\mathcal{P}$の dominate 測度という. dominate 測度をもつ$\mathcal{P}$を dominated であるという. $\mathcal{A}$がσ-generated ならば, $\mathcal{P}$が dominated であるためには$\mathcal{P}$に距離$\|P-P'\|_{\mathcal{A}}$で位相を入れたときそれが可分となることが必要十分である, ここで signed 測度$P-P'$につけられた$\|\cdot\|_{\mathcal{A}}$は$\mathcal{A}$における全変分をあらわすものとする(後に$\mathcal{A}$の部分σ代数$\mathcal{B}$について同じ記号$\|\cdot\|_{\mathcal{B}}$を用いるが, 同様に理解されたい)。特にこの場合$\mathcal{P}$の稠密可算部分集合を$\{P_n : n=1,2,\cdots\}$とすると, $\lambda_0=\sum_{n=1}^{\infty}2^{-n}\times P_n$は$\mathcal{P}$の dominate 測度である. これを pivotal 測度という. σ代数$\mathcal{B}$が$\mathcal{P}$のどの2つの要素P, P'の組に対しても十分であるとき, $\mathcal{B}$は$\mathcal{P}$に対して双十分であるという. $\mathcal{P}$が dominated であれば, $\mathcal{B}$が$\mathcal{P}$に対して双十分であることは十分であるための必要十分条件である(Halmos-Savage [4]).

上に述べた統計の概念をσ代数列に拡張して考えようというのである. $\mathcal{A}$の部分σ代数列$\{\mathcal{A}_n\}$は単調増大で$\mathcal{A}=\bigvee_{n=1}^{\infty}\mathcal{A}_n$をみたすとする. $\mathcal{B}_n$を$\mathcal{A}_n$の部分σ代数とする$(n=1, 2, \cdots)$. $\{\mathcal{B}_n\}$が$(\mathcal{P}, \{\mathcal{A}_n\})$に対して漸近十分であるというのは, $\mathcal{P}$のどの要素Pに対しても$\mathcal{A}_n$上の確率測度P_nが存在して$(n=$

1, 2, ⋯), 次の2条件をみたすことである：(1) $\mathcal{P}_n=\{P_n: P\in\mathcal{P}\}$ に対して$\mathcal{B}_n$は十分である，(2) $\forall P(\in\mathcal{P})$: $\|P_n-P\|_{\mathcal{A}_n}\to 0$ ([6]ではこの定義に誤りがあった．[7]をみよ)．列$\{\mathcal{B}_n\}$が$\mathcal{P}$の任意の2要素の組に対して漸近十分であるとき，これは漸近双十分であるという．漸近十分な$\{\mathcal{B}_n\}$はもちろん漸近双十分である．以下では漸近十分性と十分性の関係について，3節[D-5]で定義した(p)-$\overline{lim}\,\mathcal{B}_n$, (p)-$\underline{lim}\,\mathcal{B}_n$を利用して述べ，十分性と双十分性との関係と同様な関係が漸近十分性と漸近双十分性の間にあることを述べたい

3節[D-5]で定義した上極限・下極限の概念はBaseになる確率測度に依存するから，それが変るときは記号の中に入れる必要がある．(1)-$\overline{lim}_\lambda\mathcal{B}_n$, (1)-$\underline{lim}_\lambda\mathcal{B}_n$はBase measureが$\lambda$であることを示す．$\mathcal{P}$がただ2つの要素だけからなりたつとき，$\{\mathcal{B}_n\}$が$(\mathcal{P},\{\mathcal{A}_n\})$に対して漸近十分ならば，$(1)$-$\overline{lim}_\lambda\mathcal{B}_n$は$\mathcal{P}$に対して十分である，ここで$\lambda$は$\mathcal{P}$のdominate測度とする．したがって一般にdominatedな$\mathcal{P}$についても同様なことがなりたつ(Kudō [6])が，さらに$\mathcal{A}$がσ-generatedという条件を加えると，次の4つは同等である(λ_0: pivotal測度):

(a) $\{\mathcal{B}_n\}$は$(\mathcal{P},\{\mathcal{A}_n\})$に対して漸近十分，

(b) $\forall P(\in\mathcal{P})$: $e_{\lambda_0}\left(\frac{dP}{d\lambda_0},\mathcal{B}_n\right)\to 0$

(c) $\forall P(\in\mathcal{P})$: $\left\|\frac{dP}{d\lambda_0}-E^{\mathcal{B}_n}_{\lambda_0}\left[\frac{dP}{d\lambda_0}\right]\right\|_{1(\lambda_0)}\to 0$, ただし

$\|\cdot\|_{1(\lambda_0)}$ は測度空間 $(\Omega, \mathcal{A}, \lambda_0)$ 上の L^1 ノルムを示す.

(d) (1)-$\underline{\lim}_{\lambda_0} \mathcal{B}_n$ が $\mathcal{P}$ に対して十分.(Kusama [10]).

十分性の場合と同様に漸近十分性についても次のことがなりたつ(Kusama [10]): 上の定理と同じ条件のもとに, 漸近双十分性は十分性の必要十分条件である.

参考文献

[1] Becker, R.: Ensembles compacts de tribu, Z. Wahrscheinlichkeitstheorie verw. Geb. 29(1974), 229-234.

[2] Boylan. E.S.: Equiconvergence of martingale, Ann. Math. Statist. 42(1971), 552-559.

[3] Cramer, H.: Mathematical Methods of Statistics, 1946, Princeton Math. Ser.

[4] Halmos, P.R. and Savage, L.J.: Application of the Radon-Nikodym theorem to the theory of sufficient statistics, Ann. Math. Statist. 20(1949), 225-241.

[5] 岩田直子: σ-algebra の列の収束について(修士論文).

[6] Kudō, H.: On an approximation to a sufficient statistic including a concept of asymptotic sufficiency, Journ. Fac. Sci., Univ. of Tokyo, Sec. I, 17(1970), 273-290.

[7] ————: Correction to "On an approximation ··· ···". (同誌掲載予定).

[8] ————: A note on the strong convergence of σ-algebras, Ann. Prob. 2(1974), 76-83.

[9] ————: Appendix to [8]. (unpublished)

[10] Kusama, T.: On approximate sufficiency. (提出中)

[11] Loeve, M.: Probability Theory, 2nd ed., 1960, Van Nostrand.

[12] Neveu, J.: Mathematical Foundations of the Calculus of Probability, 1965. Holden-Day.

[13] ————: Note on the tightness of the metric on the set of complete sub-σ-algebras of a probability space, Ann. Math. Statist, 43(1972), 1369-1371.

[14] Nghien, D.N.: Convergence forte des esperances conditionelles et des projecteurs d'un espace de Hilbert, Ann. Inst. Henri Poincaré, Sec. B(1970), 9-13.

[15] Rogge, L.: Uniform inequalities for conditional expectations, Ann. Prob. 2(1974), 486-489.

[16] Shintani, T. and Ando, T.: Best approximants in L^1 space, Z.Wahrscheinlichkeitstheorie verw. Geb. 33(1975), 33-39.

[17] Tomkins, R.J.: On conditional medians, Ann. Prob. 3(1975), 375-379.